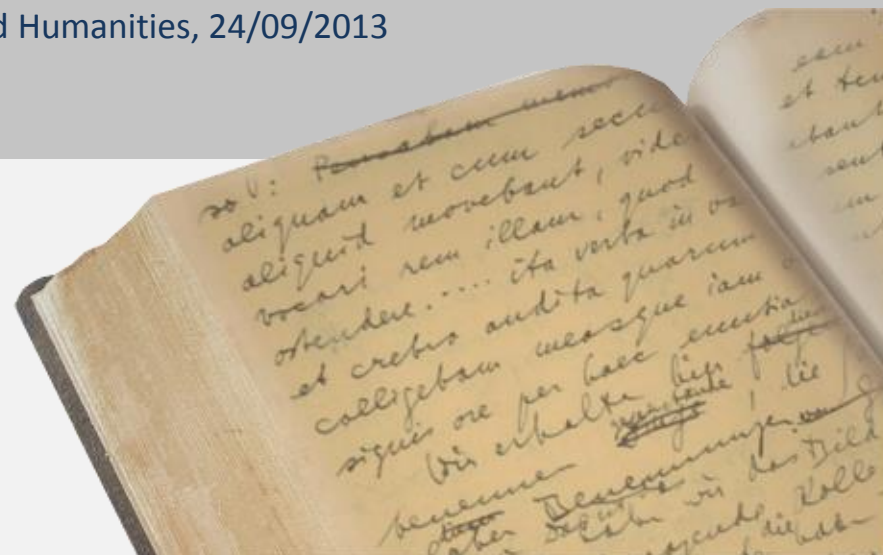
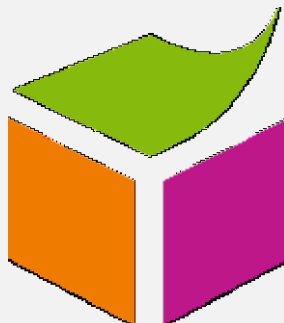


Objects, Context and Interpretation: Digitized and Digital Cultural Heritage on the Web and their Specific Potential for the Digital Humanities

Prof. Dr. Stefan Gradmann (KU Leuven)

Horizons for Social Sciences and Humanities, 24/09/2013



- Digitisation, as done to date: its benefits and limitations
- Digital Humanities and Linked Data
- DM2E and Linked Data Scholarship
- Conclusions and Recommendations with an Aside on Multilingual Operations

Digitisation as conceived today

KU LEUVEN



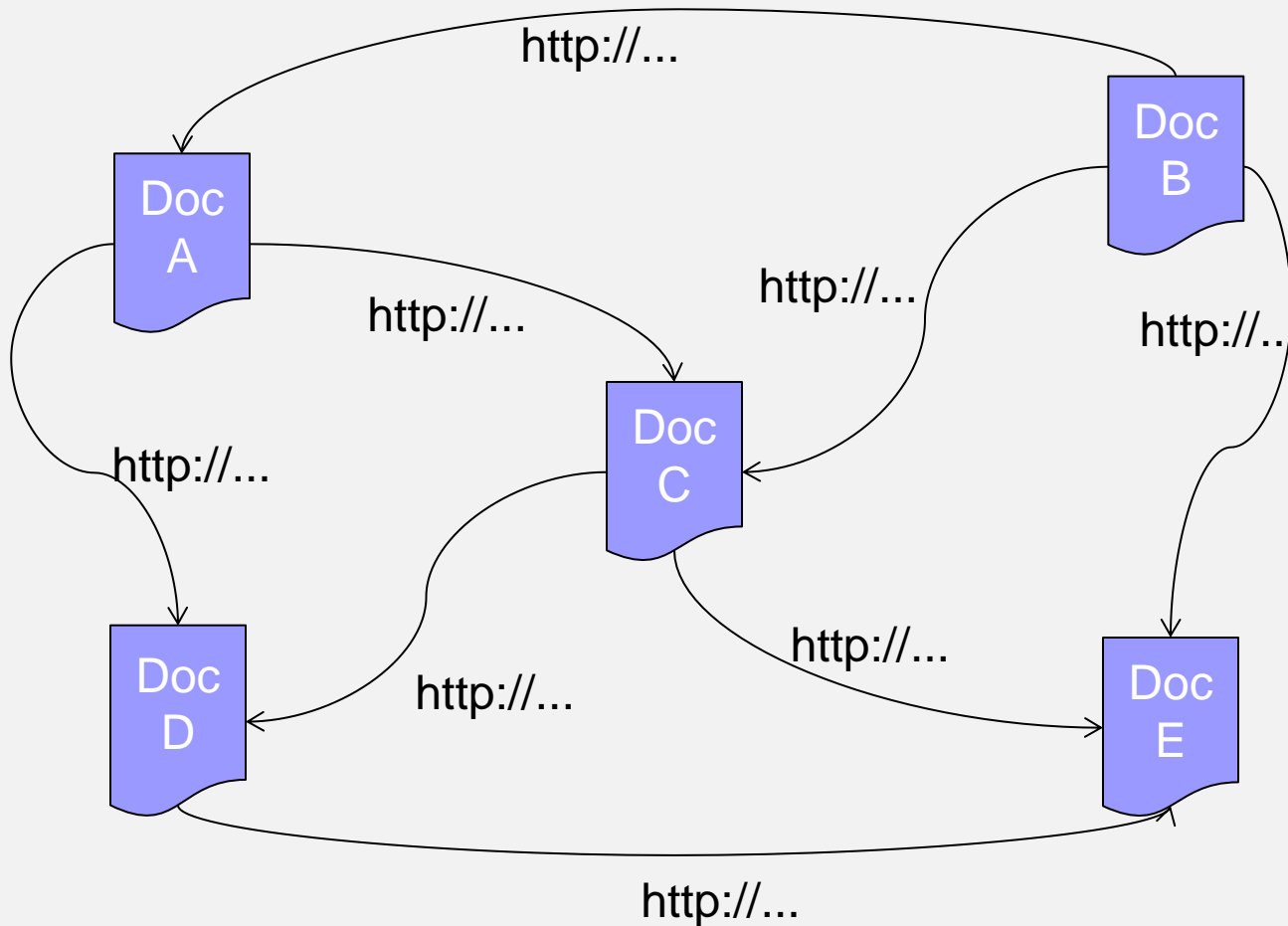
Digitized and Digital Cultural Heritage and the Humanities
Stefan Gradmann, Horizons for SSH, Vilnius, 24/09/2013



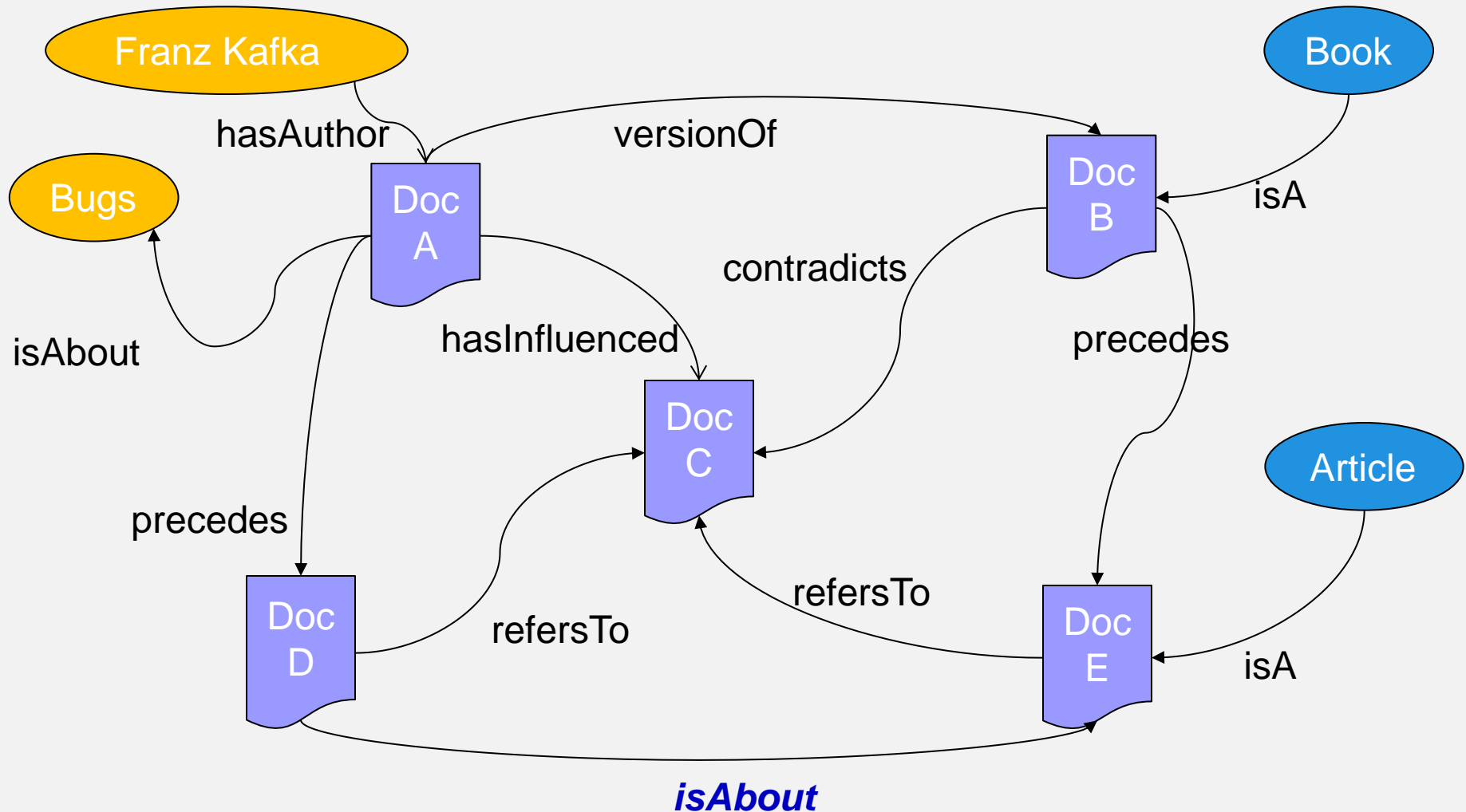
Digitisation today (2)

- Transform the **analogue** to digital,
 - focus on signal/carrier level
- **Emulates** the analogue in the digital
- Good for **preservation**
- Good for **low cost, high throughput** processing
- **Not so good for scholarly use:** “What do you do with a million books?” (G. Crane)
- **Not good at all for ‘understanding’**
- Perfectly fit for the 1st generation web

The Web of Documents ...



... extended in syntax and in scope



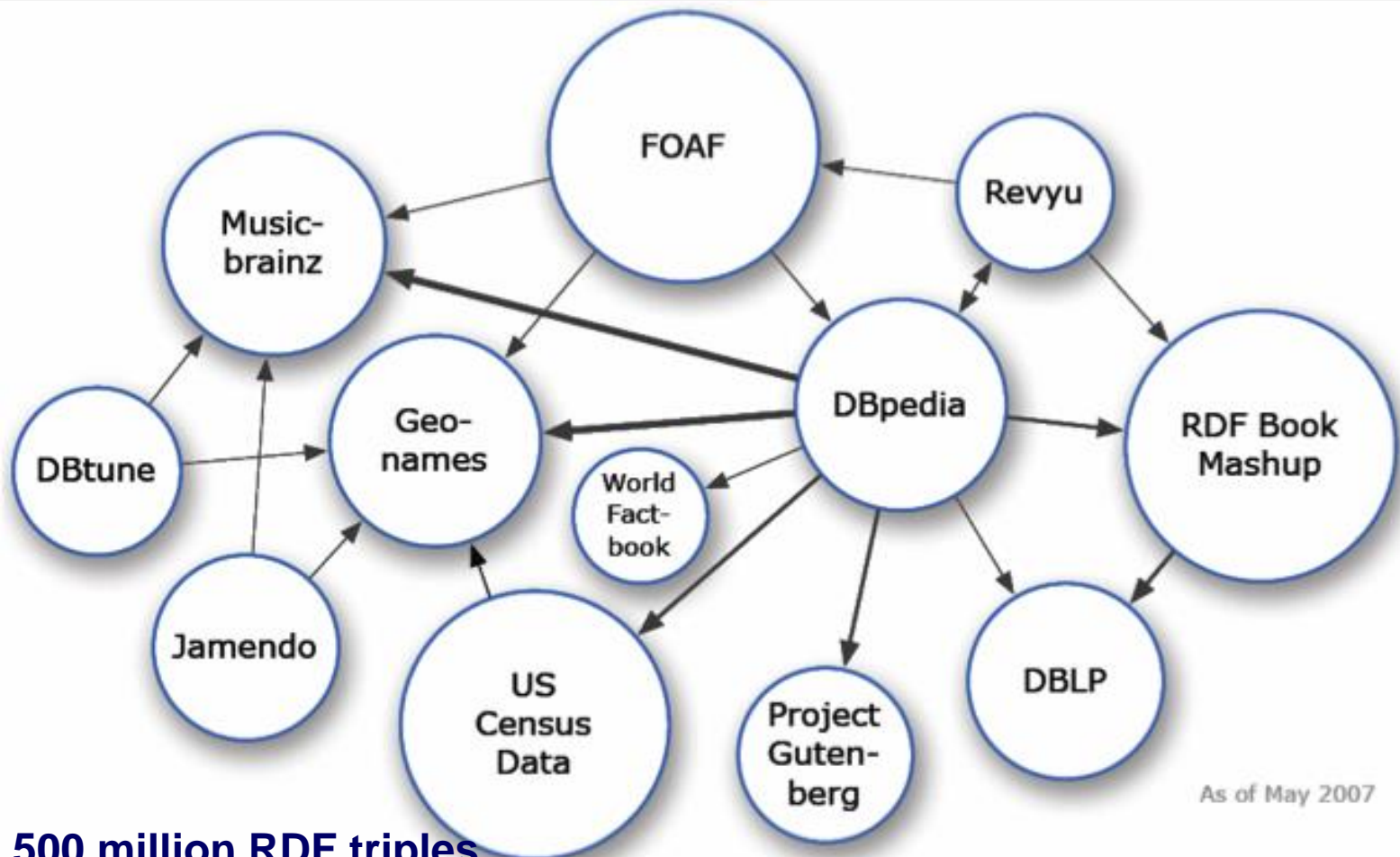
Linked Data essentials

1. Use URIs Standard **Identifiers**
 2. Use HTTP URIs Standard **Pointers**
 3. Serve useful information using SPARQL, RDF standards
 4. Mention URIs of related objects Standards for **Queries**
and **Statements**
- Link** to Context

Copyright © 2008 W3C (MIT, ERCIM, Keio)

[http://www.w3.org/2008/Talks/0617-lod-tbl/#\(4\)](http://www.w3.org/2008/Talks/0617-lod-tbl/#(4))

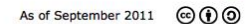
A few Bubbles: 5/2007



- Over 500 million RDF triples
- Around 120,000 RDF links between data sources

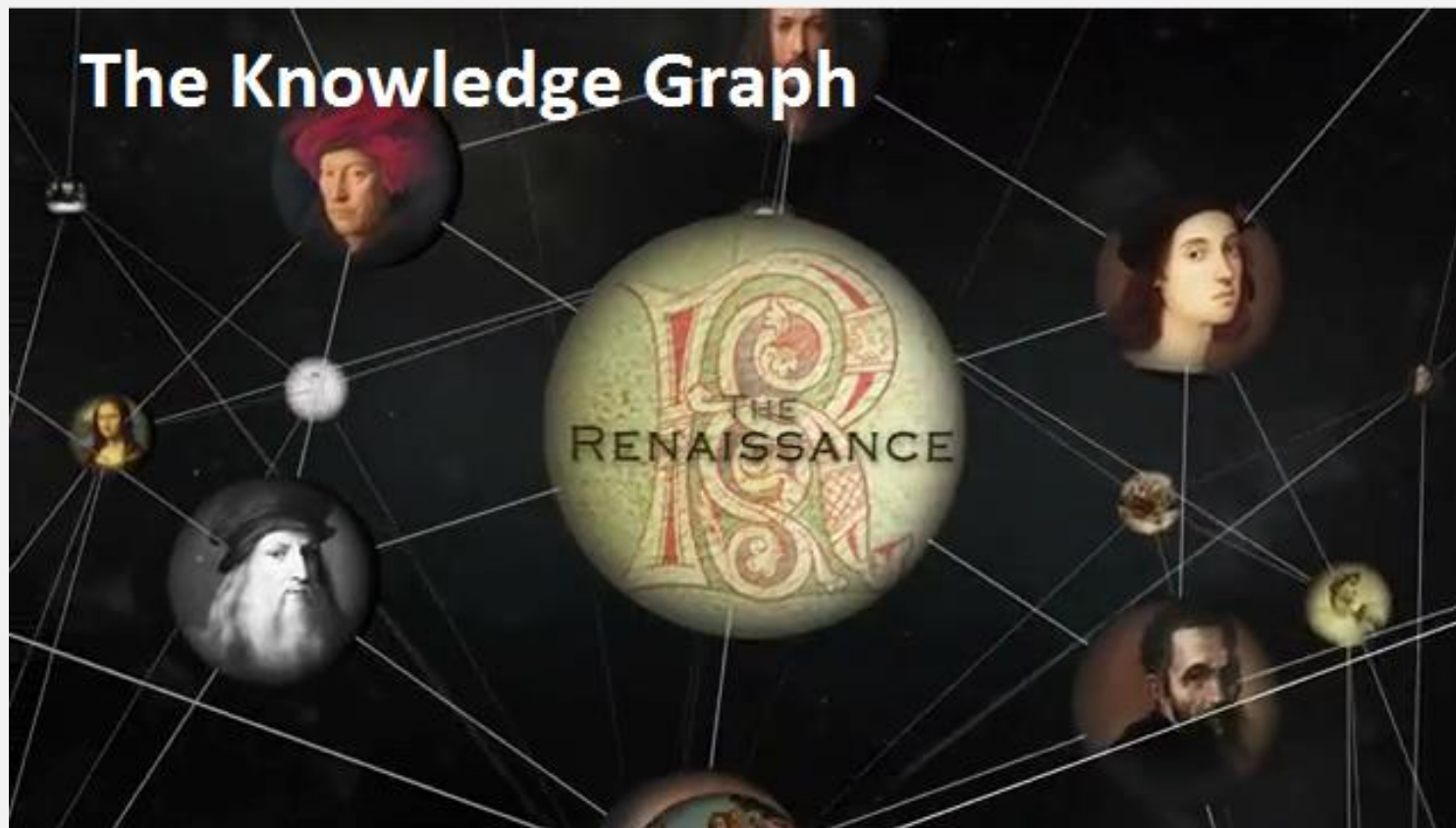
© Richard Cyganiak

KU LEUVEN



Google entering the Floor

KU LEUVEN





02/2012-01/2015

Co-funded by the EC

Digitised Manuscripts to Europeana (DM2E): What?

- Provide substantial amounts of **digital content** to Europeana with a focus on **digitised manuscripts** (WP1)
- Integrate existing **technical building blocks**
 - from Europeana development
 - as well as from generic LoD oriented development
 - into a **generic production chain** for migrating data from various sources to the EDM as well
 - as for the **contextualisation** of the object representations (WP2).
- Explore usage scenarios of EDM metadata together with object data in a specialised RDF graph based platform for **humanities research** making available specialised visualisation and reasoning environments (WP3).

WP3: Digital Humanities Requirements and Related Engineering - Context

Provide an integrated, flexible, semantic based environment targeted to **digital humanities scholars**

Structured or semi-structured data sources

Compliant web pages with metadata

Europeana (EDM)

Partners/third party platforms

Linked Open Data Cloud

Web of Data

Digital re-mix and curation tool for humanities scholars and domain experts

Text collation

Data Linking

Multimedia (semantic) annotation

Text mining

(personal) digital curation workflows

Mix, recombine digital content and metadata

Augment knowledge based on Linked Data

- Custom virtual collections
- Curated, domain specific mashups
- Scientific contributions
- etc..

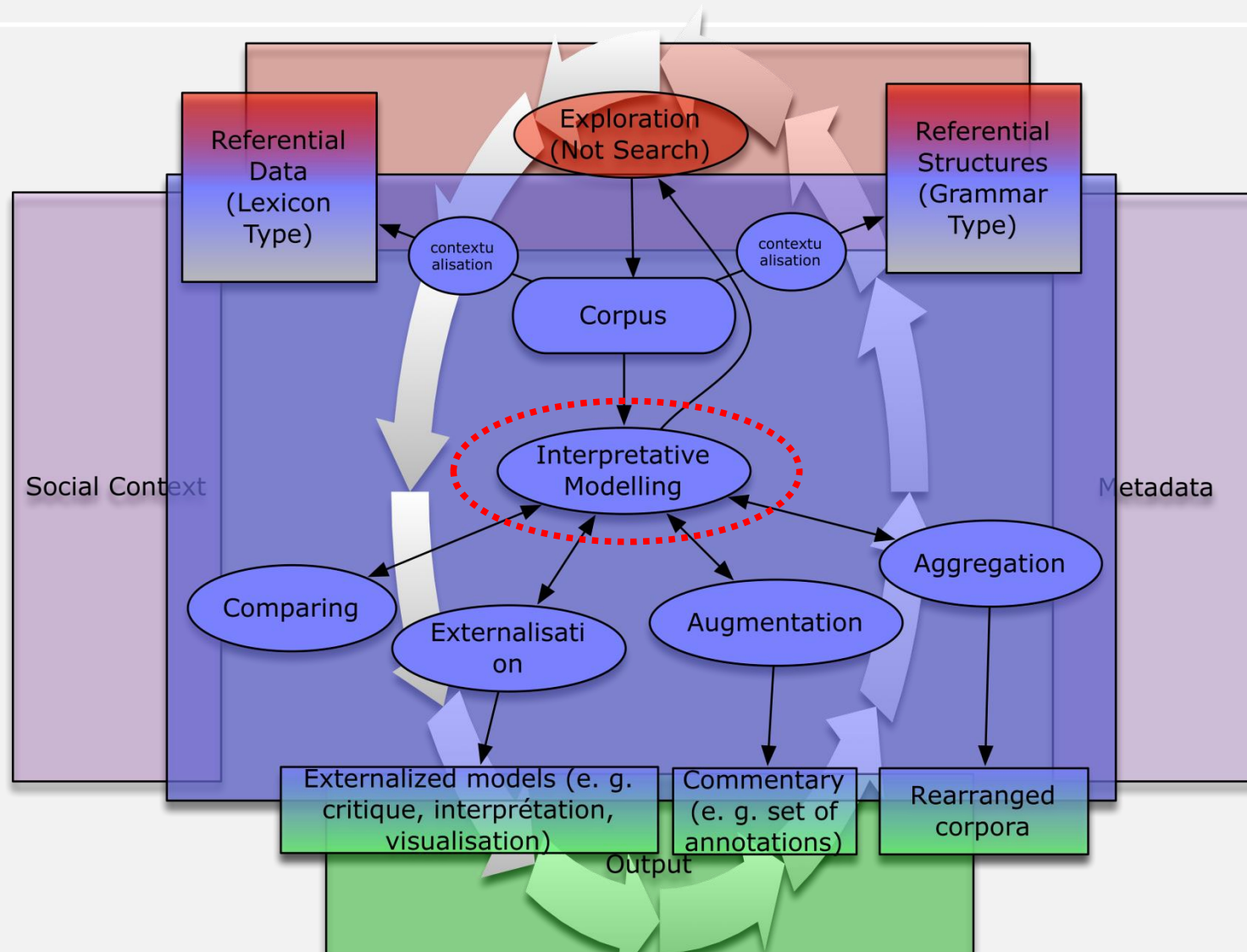
Scholarly Communities

Linked, semantically structured content further enriching the data web

- Prototype **platform enabling digital scholarship** in combining EDM RDF metadata, digitised heritage and Linked Data ...
- ... building on an **ontological representation of scholarly work...** resulting in a increasingly complex **social semantic scholarly graph** containing RDF statements such as
 - Version_A – isSuccessorOf – Version_B
 - Statement₁ – contradicts - Statement₂
 - Scribe_Y – copiedFrom – Scribe_Z

All this requires a shared understanding of the **constituent primitives** of scholarly work and of their **ontological formalisation**: concepts for modelling scholarly work and which are organised in ‘ontologies’.

A High Level Model of Research



How does this change interaction with digitised cultural heritage? The Wittgenstein Incubator

Wittgenstein SOURCE Privacy Credits

EXPLORE

DOCUMENTATION

About WittgensteinSource

SEARCH

ARCHIVE

Legend D N

Ms-114: X. Philosophische Grammatik m F D N

Ms-115: XI, Philosophische Bemerkungen m F D N

Ms-139a: Lecture on Ethics m F D N

Ms-140: Grosses Format m F D N

Ms-141 m F D N

Ms-148: C4 m F D N

Ms-149: C5 m F D N

Ms-150: C6 m F D N

Ms-152: C8 m F D N

Ms-153a: Anmerkungen m F D N

Ms-153b m F D N

Ms-154 m F D N

Ms-155 m F D N

Ms-156a m F D N

Ts-201a1: Notes on Logic m F D N

Ts-201a2: Notes on Logic m F D N

Ts-207: Lecture on Ethics m F D N

Ts-212 m F D N

Ts-213: Big Typescript m F D N

Ts-310: Brown Book m F D N

Ts-310 Facsimile

Ts-310,1 Facsimile

Ts-310,1 Ts-310,2

Ts-310,3 Ts-310,3a

Ts-310,4 Ts-310,5

Ts-310,6 Ts-310,7

WAB DISCOVERY (2008): Wittgenstein TS 310 G Trinity College Cambridge, Oxford University Press, University of Bergen

Suppose, in describing his learning of language, says that he was taught to speak by learning the names of things. It is clear that whoever says this has in mind the way in which a child learns such words as "man", "tiger", "table", etc. He does not primarily think of such words as "today", "but", "perhaps".

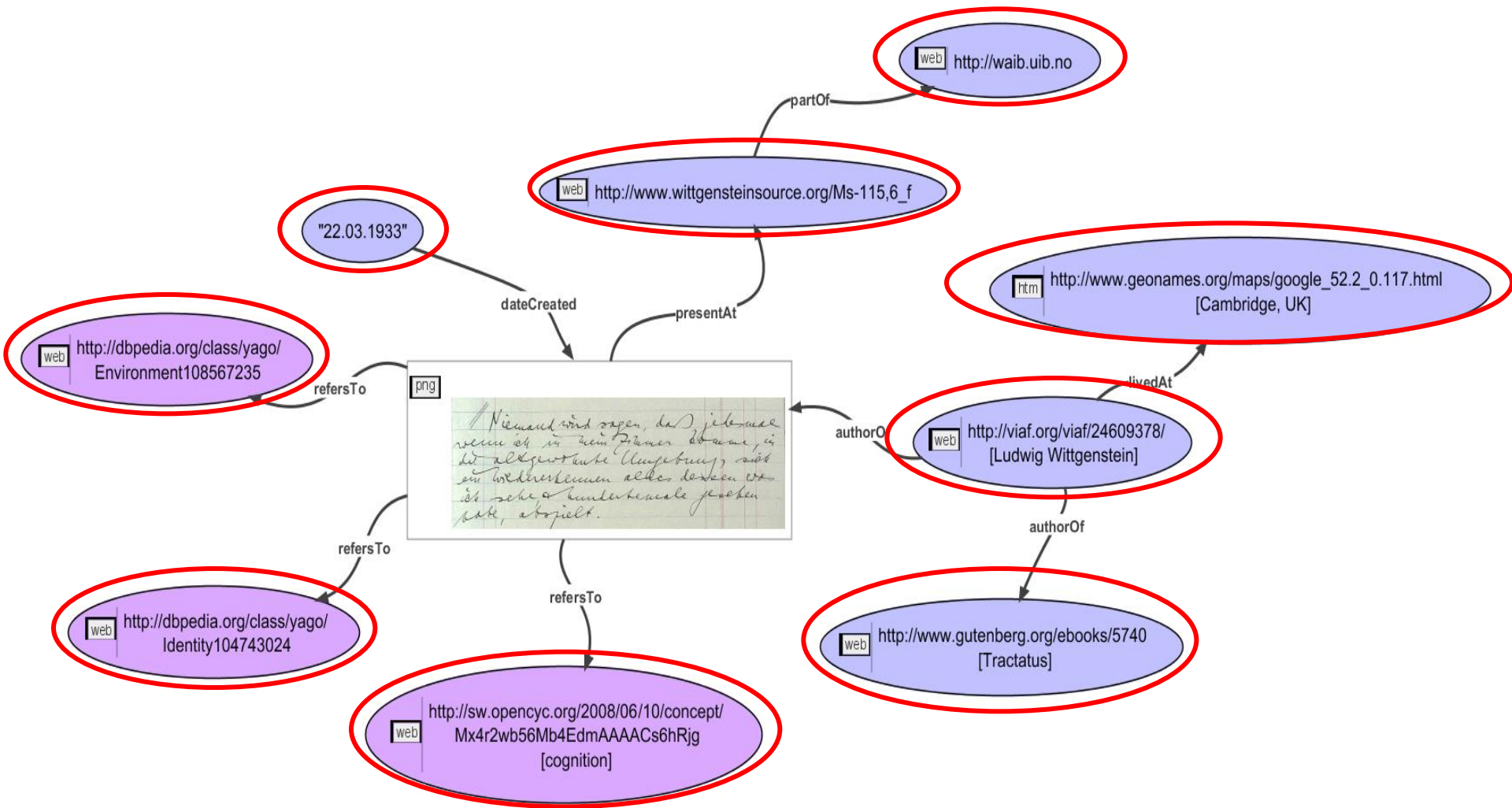
Suppose a man described a game of chess, without mentioning the existence and operation of the pawns. His description of the game as a natural phenomenon will be incomplete. On the other hand we may say that he has completely described a simpler game. In this sense we can say that Augustine's description of learning the language was correct for a simpler language than ours. Imagine this language:-

1). Its function is the communication between a builder A & his man B. B has to reach A building stones. There are cubes, bricks, slabs, beams, columns. The language consists of the words "cube", "brick", "slab", "column". A calls out one of these words, upon which B brings a stone of a certain shape. Let us imagine a society in which this is the only aspect of language. The child learns this language from the grown-up by being trained to its use. I am using the word "trained" in a way strictly analogous to that in which we talk of an animal being trained to do certain things. It is done by means of example, reward, punishment, and such like. Part of this training is that we point to a building stone, direct the attention of the child towards it, & pronounce a word. I will call this procedure ~~communicative~~ teaching of words. In the actual

Reproduced by permission of the Master and Fellows of Trinity College, Cambridge (Trinity College Library), Oxford University Press, Oxford ("Wittgenstein's Nachlass: The Bergen Electronic Edition", 2000), the University of Bergen, Bergen, and Uni Research (earlier "Unifob"), Bergen (Wittgenstein Archives at the University of Bergen). The sale, further reproduction or use of this or any image on the site for commercial purposes without prior permission from the right holders is prohibited.

Digitized and Digital Cultural Heritage and the Humanities

Contextualising Wittgenstein



Examples (Nonlinear Reading and more)

- [Graph of Thinkers](#)
- [Philosophers Edgemap](#)
- [Relfinder](#)
- [Texttexture](#)

- Much **more than just emulation of analogue object** in the digital
- Conceived on the **semantic** rather than on the **carrier** level
- Capturing rich context **enabling 'understanding'** along known and yet unknown lines of interpretation
- The result of such 'semantic' digitisation is enabling scholarly activities and interaction **that were not possible in the analog domain**
- And passing from the natural language to the concept level might be the right way to tackle the issue of **multilingual access and processing**
- This results in much more than just a substitute of 'original' digitized heritage: in this approach, digitization creates something new: **digital cultural heritage**

- Do not leave semantic digitisation to **librarians or scholars** alone: they need to work jointly on methods and ontologies
- Do not just leave digitisation to member states: **this results in awkward disparities** both in quantity and in quality as still clearly visible in the content repartition in Europeana
- The Digital Humanities need a clear, EC driven political and funding **priority for digitisation in a holistic view**
- Such a policy should have **high priority** within the first calls of Horizon 2020: this is about creating the **base layers for digital scholarship!**
- Such a policy needs to include the related activities for **ontology building and contextualisation!**

Thank you for your patience and attention!